



Docket No.: 1252.1071CIP3

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES**

In re the Application of:

Geoffrey WEHRMAN et al.

Serial No. 10/620,387

Group Art Unit: 2163

Confirmation No. 8762

Filed: July 17, 2003

Examiner: Helene Roberta Rose

For: RELOCATION OF METADATA SERVER WITH OUTSTANDING DMAPI REQUESTS

APPEAL BRIEF

Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

Sir:

I. Real Party in Interest

The inventors, Geoffrey WEHRMAN and Dean ROEHRICH, assigned all rights in the subject application to SILICON GRAPHICS, INC. according to the Assignment executed June 30, 2003, which was recorded at Reel 14303, Frames 420-422. Therefore, the real party in interest is SILICON GRAPHICS, INC.

II. Related Appeals and Interferences

There are no related appeals or interferences known to Appellants, Appellants' legal representatives or the Assignee, SILICON GRAPHICS, INC., which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

III. Status of Claims

Claims 1-12 are pending and stand rejected under 35 U.S.C. § 102(e).

IV. Status of Amendments

No Amendment was filed in response to the final Office Action mailed August 4, 2006.

V. Summary of Claimed Subject Matter

The application is directed to relocation of a metadata server managing access to a storage area network (SAN), in particular when there are outstanding requests from nodes using the data migration application programming interface (DMAPI). The relocation of the metadata server is described in paragraphs [0074] to [0082] of the application. The effect of relocation on DMAPI requests is described in paragraphs [0079] to [0081] with reference to Figs. 4A, 4B, 7, 13, 15A and 15B. As described in paragraphs [0077] and [0078], at the start of read operation 404 (Fig. 15A), the behavior head 53 (Fig. 4A) is locked 406 to prevent the list of behaviors linked off behavior head 53 from changing during I/O operations by locking the vnode or virtual metadata. Since it can take a relatively long time to resolve DMAPI requests, the read operation illustrated in Fig. 15A could be held off for an unacceptably long time during relocation of the file system being accessed. To avoid generating a migrating error that is returned to the user, behavior head 53 is unlocked during execution of `dmapi_bnc` to permit vnode 42 to be converted. A loop is executed while waiting for conversion and relocation to complete making the lock available in block 420 of Fig. 15B, so that processing of the DMAPI request can proceed.

Claim 1

Claim 1 recites a "method of executing operations on virtual metadata" (claim 1, line 1). Examples of operations performed on virtual metadata or "vnode" (represented by blocks 42a and 42b in Fig. 3 and blocks 42 in Figs. 4A and 4B) are provided in paragraph [0078] of the application and "vnode operations" are represented by blocks 58a-58d in Figs. 4A and 4B.

The body of claim 1 recites "releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata" (claim 1, lines 2-3). As described in the last sentence of paragraph [0080], "behavior head 53 is unlocked 418 during execution of `dmapi_bnc` to permit vnode 42 to be converted and the thread enters a loop waiting for conversion and relocation to complete" and the reference characters in this quotation can be found in Figs. 4A and 15B.

Claim 5

Claim 5 recites a "method of relocating a metadata server in a network of computer system nodes in which DMAPI has been implemented" (claim 5, lines 1-2). An example of this process is described in paragraphs [0074] to [0082] of the application with reference to Figs. 4A, 4B, 6, 7, 13, 14A-14H, 15A and 15B.

The body of claim 5 begins by reciting "retargeting objects on the computer system nodes accessing a current metadata server to a new metadata server" (claim 5, lines 3-4). As described in the first sentence of paragraph [0076], "the source metadata server initiates 310 retargeting and creation of client structures (objects) for the vnodes" with reference to Fig. 13.

Claim 5 also recites "releasing a lock on virtual metadata when relocation of the metadata server is underway during execution of operations on the virtual metadata" (claim 5, lines 5-6). As discussed above, the last sentence of paragraph [0080] states, "behavior head 53 is unlocked 418 during execution of dmapi_bnc to permit vnode 42 to be converted and the thread enters a loop waiting for conversion and relocation to complete" and the reference characters can be found in Figs. 4A and 15B.

Claim 9

Claim 9 recites a "cluster of computer systems" (claim 9, line 1). An example of a cluster is illustrated in Fig. 2 and described in paragraph [0021].

The body of claim 9 begins by reciting "storage devices storing at least one file" (claim 9, line 2). Examples of storage devices are disk drives 28 illustrated in Fig. 2 and described in paragraph [0021].

Claim 9 also recites "a storage area network coupled to said storage devices" (claim 9, line 3). An example of a storage area network is represented by Fibre Channel switch 30 and Fibre Channel connections 32 in Fig. 2, as described in paragraph [0021].

Claim 9 also recites "at least one metadata server node, coupled to said storage area network" (claim 9, line 4). The term "metadata server node" is used in paragraph [0043]. In the example described in paragraph [0026], node 22b is referred to as "a metadata server for the other nodes 22, 24, 26 in the cluster" (first sentence).

Claim 9 also recites "metadata client nodes, coupled to said storage area network, to release a lock on virtual metadata when relocation of said at least one metadata server is underway during execution of operations on the virtual metadata" (claim 9, last 3 lines). As described in first sentence of paragraph [0026], "nodes 22, 24, 26 ... are thus metadata clients with respect to the file system(s) for which node 22b is a metadata server." As discussed above, the last sentence of paragraph [0080] states, "behavior head 53 is unlocked 418 during execution of dmapi_bnc to permit vnode 42 to be converted and the thread enters a loop waiting for conversion and relocation to complete" and the reference characters can be found in Figs. 4A and 15B.

Claim 10

Claim 10 recites "[a]least one computer readable medium storing at least one program embodying a method of operating a cluster of computer system nodes" (claim 10, lines 1-2). As understood by one of ordinary skill in the art, the system disclosed in paragraph [0021], for example, would have programs embodied on at least one computer readable medium.

The body of claim 10 recites "releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata" (claim 10, lines 3-4). As discussed above, the last sentence of paragraph [0080] states, "behavior head 53 is unlocked 418 during execution of dmapi_bnc to permit vnode 42 to be converted and the thread enters a loop waiting for conversion and relocation to complete" and the reference characters can be found in Figs. 4A and 15B.

VI. Grounds of Rejection to be Reviewed on Appeal

As stated in the August 4, 2006 Office Action claims 1-12 were rejected under 35 USC § 102(e) as anticipated by U.S. Patent No. 6,751,616 to Chan. Therefore, at issue is whether Chan discloses all of the limitations recited in claims 1-12.

VII. Argument

As discussed in the Amendment filed May 24, 2006, Chan is directed to re-mapping of resource-to-master node assignments when the membership of a cluster changes. The issues related to changes in cluster management are discussed in the application in paragraphs [0053] to [0073] with reference to Figs. 10-12. Most of Chan is concerned with avoiding changes to the constant hash function when there is no change in the metadata server(s), only a change in the number of client nodes. A detailed description of the "re-mastering" process begins at column 16, line 6 with reference to Figs. 7A to 10B. The objective of this process is to perform "re-mastering without freezing lock operations. This improves performance of a DLM by avoiding the suspension of all lock operations during re-mastering" (column 7, lines 44- 46). Thus, when the "re-mastering" includes relocation of a master node, it appears that Chan teaches against "releasing a lock on the virtual metadata if relocation of a required metadata server is underway" (claim 1, lines 2-3). Note that in at least one embodiment, an entire metadata server is not being relocated in the process described in Chan, all that is involved in "re-mastering" is "moving some hash value ranges from one master node to another" (column 17, lines 23-24).

The process of transferring lock information to a new master is described in Chan at column 19, line 25 to column 21, line 27 with reference to Fig. 9. As noted above, the objective of this process is that "lock information is transferred to the new master in a manner that reduces or eliminates freezing out of lock requests during the transfer" (column 19, lines 26-29). The procedure illustrated in Fig. 9 and described in columns 19-21 primarily deals with how "new lock requests" (column 20, line 31) and "subsequent lock requests" (column 20, line 55) are handled. The transfer of existing locks occur in block 917 of Fig. 9 and as described at column 20, line 58 to column 21, line 2, all that happens is that "the old master node ... sends the updated master RLO for this hash value range to the new master followed by a phase IV done message" (column 20, lines 62-65). No mention has been found of "releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata" as recited in claim 1.

In the "Examiner Responses" in item 6 on page 6 of the August 4, 2006 Office Action, the Examiner cited column 5, lines 28-34 and column 11, lines 35-41 of Chan stating "metadata server is interpreted to be moving from one node to another due to administrative actions" (August 4, 2006 Office Action, page 6, lines 12-13). The cited portion of columns 5 and 11 appear to be related to what happens when "a node leaves the system, [and] the system is reconfigured to reflect the current cluster of available active nodes" (column 4, lines 20-21). The cited portion of column 11 starts with an assumption that "one message can hold the lock information for one resource being moved from an old master node to a new master node" (column 11, lines 35-37) and then describes a formula for "the total number of messages required" (column 11, lines 37-38) for "each of the n surviving nodes ... [to] be assigned as the master node for a[n] equal portion of the resources that need new masters" (column 11, lines 32-34).

While the portion of column 11 cited in the "Examiner Responses" of the August 4, 2006 Office Action indicates that relocation of masters is typically involved in the reconfiguration process disclosed by Chan, it does not indicate what Chan discloses about "relocation of a required metadata server" (e.g., claim 1, line 2), other than the number of messages required. The appealed claims are not directed to reducing the number of such messages, as Chan evidently is, but rather, what is done during relocation of a metadata server.

All of the independent claims recite operations performed during "relocation of ... metadata server" with the ellipsis (...) representing the words "a required" on line 2 of claim 1 and line 3 of claim 10, "the" on line 5 of claim 5 and "said at least one" on line 6 of claim 9. Nothing has

been cited or found in Chan that describes what is done during relocation of a metadata server, other than reconfiguration of cluster membership.

The only statements that have been found in Chan regarding "releasing a lock on the virtual metadata" (claim 1, line 2) under any circumstances are: "the requester must wait until the database server holding the granted lock releases the granted lock" (column 2, lines 58-60), "[t]he lock manager unit on the master node for the resource controls the allocation and release (or 'de-allocation') of locks for the associated resource" (column 3, lines 32-35), "global resource information ... is used for lock conversion (... changing grants to releases)" (column 9, line 66 to column 10, line 5) and "to reflect the fact that some nodes are terminating... locks granted to the processes on the terminating node(s) ... [are] released to the next lock request in the queue" (column 23, lines 29-33). However, the claims recite "releasing a lock ... if relocation of a required metadata server is underway during execution of the operations on the virtual metadata" (e.g., claim 1, lines 2-3). The first three statements in Chan regarding release of locks have no relevance to unlocking virtual metadata of a "vnode" during relocation and the last statement refers to "locks granted to the processes on the **terminating** node(s)" not to locks granted to processes on nodes that are not terminating when a metadata server node is being relocated.

In the rejection of claims 1, 5 and 10 on pages 2-3 of the August 4, 2006 Office Action, column 5, lines 56-67 of Chan was cited in support of the statement "relocating metadata server to be interpreted as data that is used to describe other data within a network environment" (August 4, 2006 Office Action, page 2, lines 21-22). However, this paragraph in Chan does not does not appear to be related to "releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata" (claim 1, lines 2-3), but rather to an activity that occurs "[i]n response to a number of open locks on each active node of the cluster" (Chan, column 5, lines 58-59). As a result of detecting a number of open locks, "a hash value range associated with a first master node is remapped to a different second master node" (column 5, lines 59-60). As discussed above, Chan defines "re-mapping" as "changing the resource-to-master node assignments" (column 4, lines 32-33). As described in Chan, this appears to usually involve a re-balancing between two master nodes, both of which continue to run in the system, as opposed to "relocation of a required metadata server" as recited in the claims.

Even if the activity described in the last paragraph of column 5 of Chan is considered equivalent to relocation of a metadata server as recited in the claims, the cause and effect is exactly the opposite of what is recited in the claims. In the claims, locks on virtual metadata are

released or opened "if relocation of a required metadata server is underway" (claim 1, lines 2-3). The exact opposite is described in the cited paragraph of Chan, i.e., if locks are open then an activity which for sake of argument might be considered equivalent to relocation of metadata server is initiated. Therefore, even if the activity described in column 5, lines 56-67 is considered equivalent to relocation of a metadata server, the operation recited at lines 2-3 of claim 1 is not described as being performed, because the locks are not released "if relocation of a required metadata server is underway" (claim 1, lines 2-3), but rather relocation is initiated if too many locks are open.

At page 3, lines 7-8 of the August 4, 2006 Office Action, column 9, line 65 to column 10, line 9 and column 12, lines 6-10 were cited as disclosing "releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata" as recited in claim 1. The paragraph spanning columns 9 and 10 of Chan starts by describing the contents of a master resource locking object (MRLO). An MRLO is described as storing "global resource information such as the resource name, the number of opened locks, the granted lock mode (shared or exclusive), a list of locks currently granted and a list of lock requests" (column 9, line 66 to column 10, line 2) that are "used for lock conversion (changing requests to grants, changing one form of granted lock to another, and changing grants to releases) and also for recovery" (column 10, lines 2-5). Recovery is defined as "a process that corrects a database, when the database server cannot complete a transaction of interdependent data manipulation operations, by returning the database to its state before the transaction began" (column 10, lines 5-9). Thus, no mention or suggestion has been found in Chan of "releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata" as recited in claim 1.

In item 7 of the "Examiner Responses" on pages 6-7 of the August 4, 2006 Office Action, column 2, lines 55-60 was cited as allegedly disclosing "*releasing or opening locks*" (August 4, 2006 Office Action, page 6, line 17, italics in original). The sentence cited in column 2 of Chan merely describes the conventional function of a lock. No mention of whether "relocation of a required metadata server is underway during execution of the operations on the virtual metadata" (claim 1, lines 2-3) appears in the cited portion of column 2. The preceding sentences in the paragraph at column 2, lines 44-60 of Chan indicate that the cited paragraph is describing what happens when "one database server requests a lock on a resource while another database server has a lock on the resource" (column 2, lines 49-50). No explanation was provided in the "Examiner Responses" of the August 4, 2006 Office Action or the Advisory Action mailed

November 21, 2006 regarding why the paragraph at column 2, lines 44-60 of Chan is relevant to "relocation of a required metadata server" (e.g., claim 1, line 3).

Claims 5 and 10 recite limitations on line 5 and lines 3-4, respectively, that are similar to the limitations in the body of claim 1. Therefore, it is submitted that claims 1, 5 and 10, as well as claims 2-4, 6-8, 11 and 12 which depend therefrom, patentably distinguish over Chan for at least the reasons discussed above and the rejection of these claims should be reversed.

Claim 9 recites a "cluster of computer systems" (line 1) that includes "metadata client nodes... to release a lock on virtual metadata when relocation of said at least one metadata server is underway during execution of operations on the virtual metadata" (claim 9, last three lines). It is submitted that claim 9 patentably distinguishes over Chan for the reasons discussed above with respect to claims 1, 5 and 10 and the rejection of claim 9 should be reversed.

Summary of Arguments

For the reasons set forth above and in the Amendment filed May 24, 2006 and the Request for Reconsideration filed November 6, 2006, it is submitted that claims 1-12 are not anticipated by Chan. Thus, it is respectfully submitted that the Examiner's final rejection of the claims is without support and, therefore, erroneous. Accordingly, the Board of Patent Appeals and Interferences is respectfully urged to so find and to reverse the Examiner's final rejection.

Please charge any required fee to our Deposit Account No. 19-3935.

Respectfully submitted,

STAAS & HALSEY LLP

Date: February 5, 2007

By: Richard A. Gollhofer
Richard A. Gollhofer
Registration No. 31,106

1201 New York Avenue, NW, Suite 700
Washington, D.C. 20005
Telephone: (202) 434-1500
Facsimile: (202) 434-1501

CERTIFICATE UNDER 37 CFR 1.8(a)
I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on 2/5/2007
STAAS & HALSEY
By: Richard A. Gollhofer
Date: 2/5/07

VIII. Claims Appendix

1. A method of executing operations on virtual metadata, comprising:
releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata.

2. A method as recited in claim 1, wherein the virtual metadata is formed as a private data chain, and said method further comprises locking a pointer to the private data chain prior to linking to a first item of private data in the private data chain.

3. A method as recited in claim 2, further comprising waiting, after said releasing, for availability of a lock on the pointer to the private data chain upon completion of relocation of the metadata server, before continuing with execution of operations on the virtual metadata.

4. A method as recited in claim 3, wherein said releasing, waiting and continuing execution of operations on the virtual metadata after relocation of the metadata server are performed transparently to users.

5. A method of relocating a metadata server in a network of computer system nodes in which DMAPI has been implemented, comprising:
retargeting objects on the computer system nodes accessing a current metadata server to a new metadata server; and
releasing a lock on virtual metadata when relocation of the metadata server is underway during execution of operations on the virtual metadata.

6. A method as recited in claim 5, wherein the virtual metadata is formed as a private data chain, and said method further comprises locking a pointer to the private data chain prior to linking to a first item of private data in the private data chain.

7. A method as recited in claim 6, further comprising waiting, after said releasing, for availability of a lock on the pointer upon completion of relocation of the metadata server, before continuing with execution of operations on the virtual metadata.

8. A method as recited in claim 7, wherein said releasing, waiting and continuing execution of operations on the virtual metadata after relocation of the metadata server are performed transparently to users.
9. A cluster of computer systems, comprising:
storage devices storing at least one file;
a storage area network coupled to said storage devices;
at least one metadata server node, coupled to said storage area network; and
metadata client nodes, coupled to said storage area network, to release a lock on virtual metadata when relocation of said at least one metadata server is underway during execution of operations on the virtual metadata.
10. At least one computer readable medium storing at least one program embodying a method of operating a cluster of computer system nodes, said method comprising:
releasing a lock on the virtual metadata if relocation of a required metadata server is underway during execution of the operations on the virtual metadata.
11. At least one computer readable medium as recited in claim 10, wherein the virtual metadata is formed as a private data chain, and said method further comprises locking a pointer to the private data chain prior to linking to a first item of private data in the private data chain.
12. At least one computer readable medium as recited in claim 11, wherein said method further comprises waiting, after said releasing, for availability of a lock on the pointer to the private data chain upon completion of relocation of the metadata server, before continuing with execution of operations on the virtual metadata.

IX. Evidence Appendix

(None)

X. Related Proceedings Appendix

(None)